STIC Biotechnology Systems Branch

RAW SEQUENCE LISTING
ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: _10/505,328A_
Source: _PCT_
Date Processed by STIC: _4/28/06_

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.
PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:
1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

FOR CRF SUBMISSION AND PATENTIN SOFTWARE QUESTIONS, PLEASE CONTACT MARK SPENCER, TELEPHONE: 571-272-2510; FAX: 571-273-0221

TO **REDUCE** ERRORED SEQUENCE LISTINGS, **PLEASE** USE THE **CHECKER VERSION 4.4.0 PROGRAM**, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS:
  http://www.uspto.gov/web/offices/pac/checker/chkrnote.htm

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail. Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom. Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses:

1. EFS-Bio (<http://www.uspto.gov/ebc/efs/downloads/documents.htm> , EFS Submission User Manual - ePAVE)
2. U.S. Postal Service: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450
3. Hand Carry, Federal Express, United Parcel Service, or other delivery service (EFFECTIVE 01/14/05): U.S. Patent and Trademark Office, Mail Stop Sequence, Customer Window, Randolph Building, 401 Dulany Street, Alexandria, VA 22314

Revised 01/10/06

PCT

RAW SEQUENCE LISTING   DATE: 04/28/2006
PATENT APPLICATION: US/10/505,328A   TIME: 09:38:57

Input Set : A:\Sequence.txt
Output Set: Ŋ:\CRF4\04282006\J505328A.raw

```
 2 <110> APPLICANT: Korea Advanced Institute of Science and Technology
 4 <120> TITLE OF INVENTION: CONSTRUCTION OF NOVEL STRAINS CONTAINING MINIMIZING
 5       GENOME BY Tn5-COUPLED Cre/loxP EXCISION SYSTEM
 7 <130> FILE REFERENCE: 02730.0020.PCUS00
 9 <140> CURRENT APPLICATION NUMBER: 10/505;328A
11 <141> CURRENT FILING DATE: 2004-08-23
11 <150> PRIOR APPLICATION NUMBER: PCT/KR02/02033
12 <151> PRIOR FILING DATE: 2002-10-31
14 <150> PRIOR APPLICATION NUMBER: KR 10-2002-0009647
15 <151> PRIOR FILING DATE: 2002-02-22
17 <160> NUMBER OF SEQ ID NOS: 13
19 <170> SOFTWARE: KopatentIn 1.71
21 <210> SEQ ID NO: 1
22 <211> LENGTH: 2437
23 <212> TYPE: DNA
24 <213> ORGANISM: Artificial Sequence
26 <220> FEATURE:
27 <223> OTHER INFORMATION: chemically synthesized TnKGloxP
30 <400> SEQUENCE: 1
31 attcaggctg cgcaactgtt gggaagggcg atcggtgcgg gcctcttcgc tattacgcca         60
33 gctgtctctt atacacatct caaccatcat cgatgaattc gagctcggta cccgggttga        120
35 actgcggatc ttgcggccgc aaaaattaaa aatgaagttt tgacggtatc gaaccccaga        180
37 gtcccgctca gaagaactcg tcaagaaggc gatagaaggc gatgcgctgc gaatcgggag        240
39 cggcgatacc gtaaagcacg aggaagcggt cagcccattc gccgccaagc tcttcagcaa        300
41 tatcacgggt agccaacgct atgtcctgat agcggtccgc cacacccagc cggccacagt        360
43 cgatgaatcc agaaaagcgg ccattttcca ccatgatatt cggcaagcag gcatcgccat        420
45 gggtcacgac gagatcctcg ccgtcgggca tccgcgcctt gagcctggcg aacagttcgg        480
47 ctggcgcgag cccctgatgc tcttcgtcca gatcatcctg atcgacaaga ccggcttcca        540
49 tccgagtacg tgctcgctcg atgcgatgtt tcgcttggtg gtcgaatggg caggtagccg        600
51 gatcaagcgt atgcagccgc cgcattgcat cagccatgat ggatactttc tcggcaggag        660
53 caaggtgaga tgacaggaga tcctgccccg gcacttcgcc caatagcagc cagtcccttc        720
55 ccgcttcagt gacaacgtcg agcacagctg cgcaaggaac gcccgtcgtg ccagccacg         780
57 atagccgcgc tgcctcgtct tggagttcat tcagggcacc ggacaggtcg gtcttgacaa        840
59 aaagaaccgg gcgcccctgc gctgacagcc ggaacacggc ggcatcagag cagccgattg        900
61 tctgttgtgc ccagtcatag ccgaatagcc tctccaccca agcggccgga gaacctgcgt        960
63 gcaatccatc ttgttcaatc atgcgaaacg atcctcatcc tgtctcttga tccactagat       1020
65 tattgaagca tttatcaggg ttattgtctc atgagcggat acatatttga atgtatttag       1080
67 aaaaataaac aaatagggggt tccgcgcaca tttccccgaa aagtgccacc tgcatcgatg       1140
69 aattgatccg aagttcctat tctctagaaa gtataggaac ttcgaattgt cgacaagctt       1200
71 gatctggctt atcgaaatta atacgactca ctataggag accggaattc attatttgta        1260
73 gagctcatcc atgccatgtg taatcccagc agcagttaca aactcaagaa ggaccatgtg       1320
75 gtcacgcttt tcgttgggat ctttcgaaag ggcagattgt gtcgacaggt aatggttgtc       1380
77 tggtaaaagg acagggccat cgccaattgg agtattttgt tgataatggt ctgctagttg      1440
```

RAW SEQUENCE LISTING   DATE: 04/28/2006
PATENT APPLICATION: US/10/505,328A   TIME: 09:38:57

Input Set : A:\Sequence.txt
Output Set: N:\CRF4\04282006\J505328A.raw

```
 79 aacggatcca tcttcaatgt tgtggcgaat tttgaagtta gctttgattc cattcttttg      1500
 81 tttgtctgcc gtgatgtata cattgtgtga gttatagttg tactcgagtt tgtgtccgag      1560
 83 aatgtttcca tcttctttaa aatcaatacc ttttaactcg atacgattaa caagggtatc      1620
 85 accttcaaac ttgacttcag cacgcgtctt gtagttcccg tcatctttga aagatatagt      1680
 87 gcgttcctgt acataacctt cgggcatggc actcttgaaa aagtcatgcc gtttcatatg      1740
 89 atccggataa cgggaaaagc attgaacacc ataagagaaa gtagtgacaa gtgttggcca      1800
 91 tggaacaggt agttttccag tagtgcaaat aaatttaagg gtaagttttc cgtatgttgc      1860
 93 atcaccttca ccctctccac tgacagaaaa tttgtgccca ttaacatcac catctaattc      1920
 95 aacaagaatt gggacaactc cagtgaaaag ttcttctcct ttactcattt tttctaccgg      1980
 97 tacccgggga tcctctagag tcgacctgca ggcatgcaag cttggcgtaa tcatggtcat      2040
 99 agctgtttcc tgtgtgaaat tgttatccgc tcacaattcc acacaacata cgagccggaa      2100
101 gcataaagtg taaagcctgg ggtgcctaat gagtgagcta actcacatta attgcgttgc      2160
103 gctcactgcc cgctttccag tcgggaaatc caagggcgaa ttcgagctcg gtaccgggcc      2220
105 cccctcgag ggacctaata acttcgtata gcatacatta tacgaagtta tattaagggt      2280
107 tccggatcct ctagagtaga cctctagagt cgacctgcag gcatgcaagc ttcagggttg      2340
109 agatgtgtat aagagacagc tgcattaatg aatcggccaa cgcgcgggga gaggcggttt      2400
111 gcgtattggg cgctcttccg cttcctcgct cactgac                              2437
114 <210> SEQ ID NO: 2
115 <211> LENGTH: 1511
116 <212> TYPE: DNA
117 <213> ORGANISM: Artificial Sequence
119 <220> FEATURE:
120 <223> OTHER INFORMATION: chemically synthesized TnCloxP
123 <400> SEQUENCE: 2
124 attcaggctg cgcaactgtt gggaagggcg atcggtgcgg gcctcttcgc tattacgcca        60
126 gctgtctctt atacacatct caaccatcat cgatgaattc gagctcggta ccgcaaaaat       120
128 taaaaatgaa gtttaaaatc aatctaaagt atatatgagt aaacttggtc tgacagttac       180
130 caatgcttaa tcagtgaggc accaataact gccttaaaaa aattacgccc cgccctgcca       240
132 ctcatcgcag tactgttgta attcattaag cattctgccg acatggaagc catcacagac       300
134 ggcatgatga acctgaatcg ccagcggcat cagcaccttg tcgccttgcg tataatattt       360
136 gcccatggtg aaaacggggg cgaagaagtt gtccatattg ccacgtttaa atcaaaact        420
138 ggtgaaactc acccagggat tggctgagac gaaaaacata ttctcaataa acccttttagg      480
140 gaaataggcc aggttttcac cgtaacacgc cacatcttgc gaatatatgt gtagaaactg       540
142 ccggaaatcg tcgtggtatt cactccagag cgatgaaaac gtttcagttt gctcatggaa       600
144 aacggtgtaa caagggtgaa cactatccca tatcaccagc tcaccgtctt tcattgccat       660
146 acggaatttc ggatgagcat tcatcaggcg ggcaagaatg tgaataaagg ccggataaaa       720
148 cttgtgctta ttttttcttta cggtctttaa aaaggccgta atatccagct gaacggtctg      780
150 gttataggta cattgagcaa ctgactgaaa tgcctcaaaa tgttctttac gatgccattg       840
152 ggatatatca acggtggtat atccagtgat ttttttctcc atttttagctt ccttagctcc      900
154 tgaaaatctc gataactcaa aaaatacgcc cggtagtgat cttatttcat tatggtgaaa       960
156 gttggaacct cttacgtgcc gatcaacgtc tcattttcgc caaaagttgg cccagggctt      1020
158 cccggtatca acagggacac caggatttat ttattctgcg aagtgatctt ccgtcacagg      1080
160 tatttattcg gcgcaaagtg cgtcgggtga tgctgccaac ttactgattt agtgtatgat      1140
162 ggtgtttttg aggtgctcca gtggcttctg tttctatcag catcgatgaa ttgatccgaa      1200
164 gttcctattc tctagaaagt ataggaactt cgaattgtcg acaagcttga tctggcttat      1260
166 cgaaattaat acgactcact atagggagac cggaattcga gctcggtacc gggcccccc       1320
168 tcgagggacc taataacttc gtatagcata cattatacga gttatatta agatcctcta       1380
170 gagtcgacct gcaggcatgc aagcttcagg gttgagatgt gtataagaga cagctgcatt      1440
172 aatgaatcgg ccaacgcgcg gggagaggcg gtttgcgtat gggcgctct tccgcttcct       1500
```

RAW SEQUENCE LISTING   DATE: 04/28/2006
PATENT APPLICATION: US/10/505,328A   TIME: 09:38:57

                Input Set : A:\Sequence.txt
                Output Set: N:\CRF4\04282006\J505328A.raw

```
174 cgctcactga c                                                          1511
177 <210> SEQ ID NO: 3
178 <211> LENGTH: 19
179 <212> TYPE: DNA
180 <213> ORGANISM: Artificial Sequence
182 <220> FEATURE:
183 <223> OTHER INFORMATION: chemically synthesized OE sequence
186 <400> SEQUENCE: 3
187 ctgtctctta tacacatct                                                    19
190 <210> SEQ ID NO: 4
191 <211> LENGTH: 34
192 <212> TYPE: DNA
193 <213> ORGANISM: Artificial Sequence
195 <220> FEATURE:
196 <223> OTHER INFORMATION: chemically synthesized loxP site
199 <400> SEQUENCE: 4
200 ataacttcgt atagcataca ttatacgaag ttat                                   34
203 <210> SEQ ID NO: 5
204 <211> LENGTH: 996
205 <212> TYPE: DNA
206 <213> ORGANISM: Artificial Sequence
208 <220> FEATURE:
209 <223> OTHER INFORMATION: chemically synthesized KmR gene
212 <400> SEQUENCE: 5
213 gcaaaaatta aaaatgaagt tttgacggta tcgaacccca gagtcccgct cagaagaact        60
215 cgtcaagaag gcgatagaag gcgatgcgct gcgaatcggg agcggcgata ccgtaaagca       120
217 cgaggaagcg gtcagcccat tcgccgccaa gctcttcagc aatatcacgg gtagccaacg       180
219 ctatgtcctg atagcggtcc gccacaccca gccggccaca gtcgatgaat ccagaaaagc       240
221 ggccattttc caccatgata ttcggcaagc aggcatcgcc atgggtcacg acgagatcct       300
223 cgccgtcggg catccgcgcc ttgagcctgg cgaacagttc ggctggcgcg agccctgat        360
225 gctcttcgtc cagatcatcc tgatcgacaa gaccggcttc catccgagta cgtgctcgct       420
227 cgatgcgatg tttcgcttgg tggtcgaatg ggcaggtagc cggatcaagc gtatgcagcc       480
229 gccgcattgc atcagccatg atggatactt tctcggcagg agcaaggtga gatgacagga       540
231 gatcctgccc cggcacttcg cccaatagca gccagtccct tcccgcttca gtgacaacgt       600
233 cgagcacagc tgcgcaagga acgcccgtcg tggccagcca cgatagccgc gctgcctcgt       660
235 cttggagttc attcagggca ccggacaggt cggtcttgac aaaaagaacc gggcgcccct       720
237 gcgctgacag ccggaacacg gcggcatcag agcagccgat tgtctgttgt gcccagtcat       780
239 agccgaatag cctctccacc caagcggccg gagaacctgc gtgcaatcca tcttgttcaa       840
241 tcatgcgaaa cgatcctcat cctgtctctt gatccactag attattgaag catttatcag       900
243 ggttattgtc tcatgagcgg atacatattt gaatgtattt agaaaaataa acaaataggg       960
245 gttccgcgca catttccccg aaaagtgcca cctgca                                 996
248 <210> SEQ ID NO: 6
249 <211> LENGTH: 947
250 <212> TYPE: DNA
251 <213> ORGANISM: Artificial Sequence
253 <220> FEATURE:
254 <223> OTHER INFORMATION: chemically synthesized GFP gene
257 <400> SEQUENCE: 6
258 attatttgta gagctcatcc atgccatgtg taatcccagc agcagttaca aactcaagaa        60
```

RAW SEQUENCE LISTING   DATE: 04/28/2006
PATENT APPLICATION: US/10/505,328A   TIME: 09:38:57

Input Set : A:\Sequence.txt
Output Set: N:\CRF4\04282006\J505328A.raw

```
260 ggaccatgtg gtcacgcttt tcgttgggat ctttcgaaag ggcagattgt gtcgacaggt    120
262 aatggttgtc tggtaaaagg acagggccat cgccaattgg agtattttgt tgataatggt    180
264 ctgctagttg aacggatcca tcttcaatgt tgtggcgaat tttgaagtta gctttgattc    240
266 cattcttttg tttgtctgcc gtgatgtata cattgtgtga gttatagttg tactcgagtt    300
268 tgtgtccgag aatgtttcca tcttctttaa aatcaatacc ttttaactcg atacgattaa    360
270 caagggtatc accttcaaac ttgacttcag cacgcgtctt gtagttcccg tcatctttga    420
272 aagatatagt gcgttcctgt acataacctt cgggcatggc actcttgaaa aagtcatgcc    480
274 gtttcatatg atccggataa cgggaaaagc attgaacacc ataagagaaa gtagtgacaa    540
276 gtgttggcca tggaacaggt agttttccag tagtgcaaat aaatttaagg gtaagttttc    600
278 cgtatgttgc atcaccttca ccctctccac tgacagaaaa tttgtgccca ttaacatcac    660
280 catctaattc aacaagaatt gggacaactc cagtgaaaag ttcttctcct ttactcattt    720
282 tttctaccgg tacccgggga tcctctagag tcgacctgca ggcatgcaag cttggcgtaa    780
284 tcatggtcat agctgtttcc tgtgtgaaat tgttatccgc tcacaattcc acacaacata    840
286 cgagccggaa gcataaagtg taaagcctgg ggtgcctaat gagtgagcta actcacatta    900
288 attgcgttgc gctcactgcc cgctttccag tcgggaaatc caagggc                  947
291 <210> SEQ ID NO: 7
292 <211> LENGTH: 1069
293 <212> TYPE: DNA
294 <213> ORGANISM: Artificial Sequence
296 <220> FEATURE:
297 <223> OTHER INFORMATION: chemically synthesized CmR gene
300 <400> SEQUENCE: 7
301 gcaaaaatta aaaatgaagt tttaaatcaa tctaaagtat atatgagtaa acttggtctg     60
303 acagttacca atgcttaatc agtgaggcac caataactgc cttaaaaaaa ttacgccccg    120
305 ccctgccact catcgcagta ctgttgtaat tcattaagca ttctgccgac atggaagcca    180
307 tcacagacgg catgatgaac ctgaatcgcc agcggcatca gcaccttgtc gccttgcgta    240
309 taatatttgc ccatggtgaa aacgggggcg aagaagttgt ccatattggc cacgtttaaa    300
311 tcaaaactgg tgaaactcac ccagggattg gctgagacga aaaacatatt ctcaataaac    360
313 cctttaggga ataggccag gtttttcaccg taacacgcca catcttgcga atatatgtgt    420
315 agaaactgcc ggaaatcgtc gtggtattca ctccagagcg atgaaaacgt ttcagtttgc    480
317 tcatggaaaa cggtgtaaca agggtgaaca ctatcccata tcaccagctc accgtctttc    540
319 attgccatac ggaatttcgg atgagcattc atcaggcggg caagaatgtg aataaaggcc    600
321 ggataaaact tgtgcttatt tttctttacg gtctttaaaa aggccgtaat atccagctga    660
323 acggtctggt tataggtaca ttgagcaact gactgaaatg cctcaaaatg ttctttacga    720
325 tgccattggg atatatcaac ggtggtatat ccagtgattt ttttctccat tttagcttcc    780
327 ttagctcctg aaaatctcga taactcaaaa aatacgcccg gtagtgatct tatttcatta    840
329 tggtgaaagt tggaacctct tacgtgccga tcaacgtctc attttcgcca aaagttggcc    900
331 cagggcttcc cggtatcaac agggacacca ggatttattt attctgcgaa gtgatcttcc    960
333 gtcacaggta tttattcggc gcaaagtgcg tcgggtgatg ctgccaactt actgatttag   1020
335 tgtatgatgg tgttttttgag gtgctccagt ggcttctgtt tctatcagc              1069
338 <210> SEQ ID NO: 8
339 <211> LENGTH: 19
340 <212> TYPE: DNA
341 <213> ORGANISM: Artificial Sequence
343 <220> FEATURE:
344 <223> OTHER INFORMATION: chemically synthesized primer-pMODFP-1
347 <400> SEQUENCE: 8
348 attcaggctg cgcaactgt                                                  19
351 <210> SEQ ID NO: 9
```

```
352 <211> LENGTH: 22
353 <212> TYPE: DNA
354 <213> ORGANISM: Artificial Sequence
356 <220> FEATURE:
357 <223> OTHER INFORMATION: chemically synthesized primer-pMODRP-1
360 <400> SEQUENCE: 9
361 tcagtgagcg aggaagcgga ag                                        22
364 <210> SEQ ID NO: 10
365 <211> LENGTH: 28
366 <212> TYPE: DNA
367 <213> ORGANISM: Artificial Sequence
369 <220> FEATURE:
370 <223> OTHER INFORMATION: chemically synthesized primer-Tn5Ext
373 <400> SEQUENCE: 10
374 agcatacatt atacgaagtt atattaag                                  28
377 <210> SEQ ID NO: 11
378 <211> LENGTH: 35
379 <212> TYPE: DNA
380 <213> ORGANISM: Artificial Sequence
382 <220> FEATURE:
383 <223> OTHER INFORMATION: chemically synthesized primer-Arb1
386 <400> SEQUENCE: 11
387 ttgagcgata gacgtacgat nnnnnnnnnn gatat                          35
390 <210> SEQ ID NO: 12
391 <211> LENGTH: 20
392 <212> TYPE: DNA
393 <213> ORGANISM: Artificial Sequence
395 <220> FEATURE:
396 <223> OTHER INFORMATION: chemically synthesized primer-Arb2
399 <400> SEQUENCE: 12
400 ttgagcgata gacgtacgat                                           20
403 <210> SEQ ID NO: 13
404 <211> LENGTH: 25
405 <212> TYPE: DNA
406 <213> ORGANISM: Artificial Sequence
408 <220> FEATURE:
409 <223> OTHER INFORMATION: chemically synthesized primer-Tn5Int
412 <400> SEQUENCE: 13
413 tcgacctgca ggcatgcaag cttca                                     25
```

W--> 387 [handwritten annotations: "pls explain "N" locations." "See error explanation on page 7."]

RAW SEQUENCE LISTING ERROR SUMMARY   DATE: 04/28/2006
PATENT APPLICATION: US/10/505,328A   TIME: 09:38:58

Input Set : A:\Sequence.txt
Output Set: N:\CRF4\04282006\J505328A.raw

Please Note:
Use of n and/or Xaa have been detected in the Sequence Listing. Please review the
Sequence Listing to ensure that a corresponding explanation is presented in the <220>
to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:11; N Pos. 21,22,23,24,25,26,27,28,29,30

VARIABLE LOCATION SUMMARY   DATE: 04/28/2006
PATENT APPLICATION: US/10/505,328A   TIME: 09:38:58

Input Set : A:\Sequence.txt
Output Set: N:\CRF4\04282006\J505328A\raw

Error Explanation: 2

Use of n's or Xaa's(NEW RULES):
Use of n's and/or Xaa's have been detected in the Sequence Listing.
Use of <220> to <223> is MANDATORY if n's or Xaa's are present.
in <220> to <223> section, please explain location of n or Xaa, and which
residue n or Xaa represents.

Seq#:11; N Pos. 21,22,23,24,25,26,27,28,29,30

VERIFICATION SUMMARY   DATE: 04/28/2006
PATENT APPLICATION: US/10/505,328A   TIME: 09:38:58

Input Set : A:\Sequence.txt
Output Set: N:\CRF4\04282006\J505328A.raw

L:11 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:387 M:258 W: Mandatory Feature missing, <221> Tag not found for SEQ ID#:11
L:387 M:258 W: Mandatory Feature missing, <222> Tag not found for SEQ ID#:11
L:387 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:11 after pos.:0